

Macintosh users: SLIS 5080 – Calculating Pearson's r (with Goofy)



As forgetful as some of us are (This certainly includes me.) we need notes and examples to refer to back to.

This came from Module 3, Part II, Slide 4.3 where we were told to refer to the Academic Library Use Study (page 11 of our text) and use the **Length of Time of Visit** and **Number of Materials Checked Out** and come up with a **Pearson's r**.

I used the following website: Free Statistics and Forecasting Software
http://www.wessa.net/rwasp_correlation.wasp#output

I will need to do multiple snapshots below, because I could not capture all this in one snapshot.

Pearson Correlation – Free Statistics and Forecasting Software (Calculators) v.1.1.23-r4

http://www.wessa.net/rwasp_correlation.wasp#output

CiteULike UNTeCampus Scirus harris.co.p.l houston.p.l my.unt.edu ecamp.unt.edu library.unt.edu

Pearson Correlation – Free Statisti...

Enter (or paste) your data delimited by hard returns.

Send output to:

Browser Blue - Charts White

Data X:

120
30
10
5
10
20
30
30
45
10
45
15
5
100
90
20
5
10
30
15

Data Y:

4
2
1
0
2
8
4
2
0
2
4
1
0
6
4
1
0
4
7
3

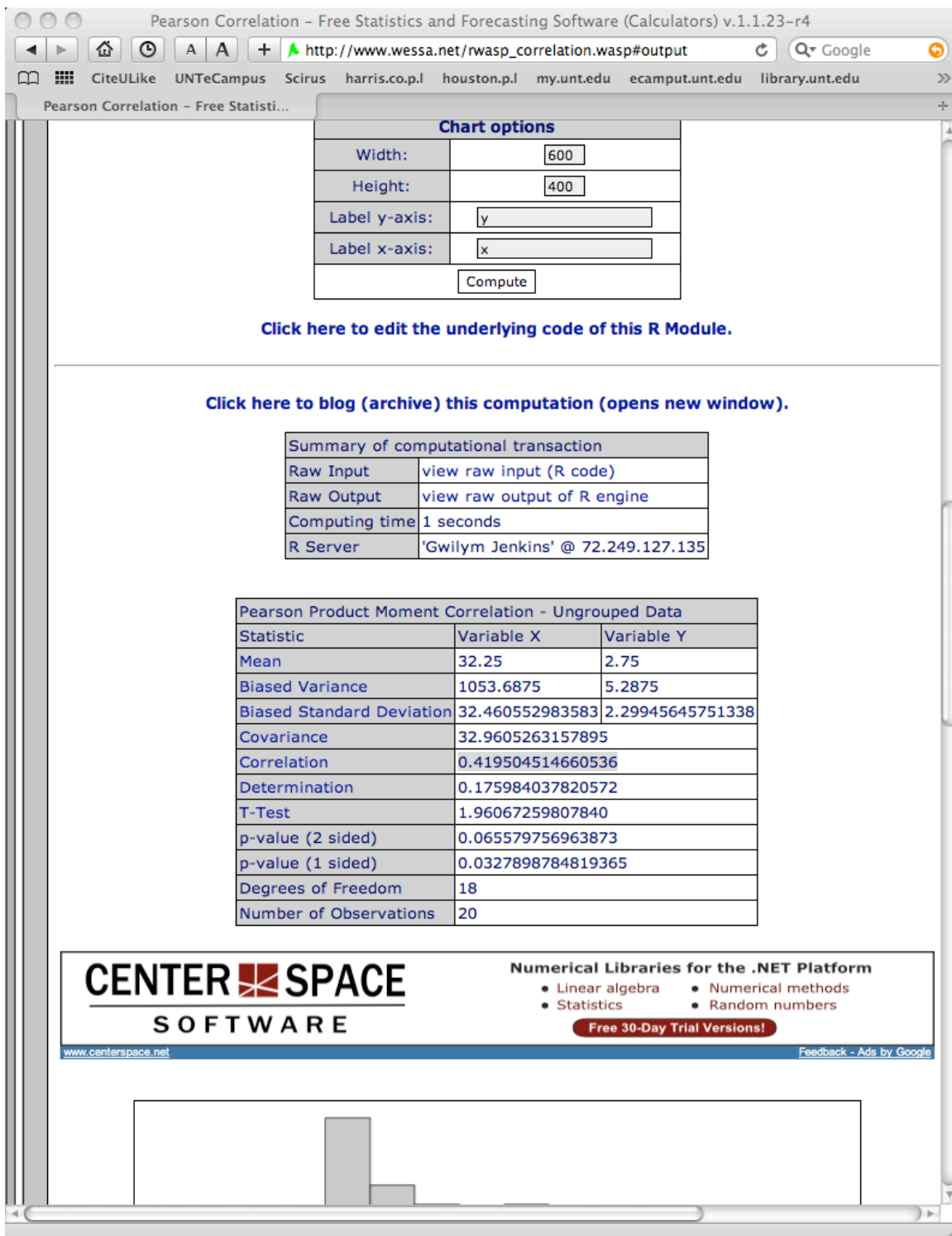
Sample Range:
(leave blank to include all observations)

From:

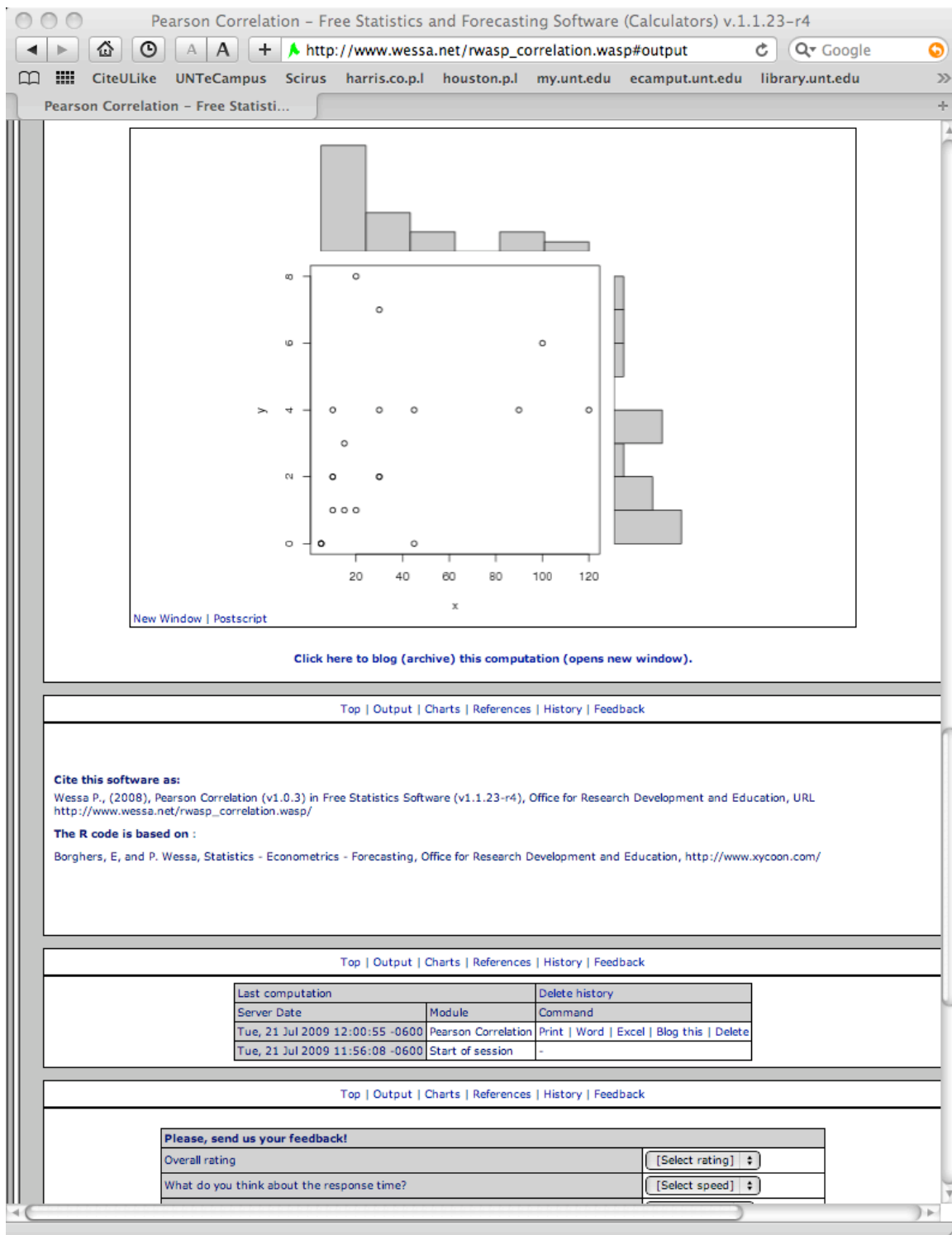
To:

Chart options

Notice how I inserted the data-X and the date-Y above.



Notice (above) the “**Correlation**” of “**0.419504514660536**” which is the Pearson’s r . It is positive, which means that when the independent variable goes up so does the dependent variable. The chart continues below:



Some of the data were repeated numbers and of course the dots on the charts are in the same spot. They appear to be a bit darker, but I really cannot tell for sure.

The data (above) **came from page 11 of our text.**

X= **Length of Time of Visit (in Minutes)**

Y=**# of Materials Checked Out**

Academic Library Use Study Data
(from page 11 of text)

Subject #	Race	Sex	Class	# of Materials Checked Out	Length of Time of Visit (in Minutes)	% of Time on Task
1	O	F	G	4	120	90
2	B	F	G	2	30	85
3	W	M	U	1	10	80
4	W	F	U	0	5	0
5	O	F	U	2	10	90
6	O	M	G	8	20	100
7	W	M	U	4	30	75
8	W	M	G	2	30	80
9	W	F	G	0	45	75
10	B	M	U	2	10	80
11	W	F	U	4	45	70
12	W	M	U	1	15	50
13	O	M	U	0	5	0
14	W	F	G	6	100	90
15	B	F	G	4	90	95
16	W	F	U	1	20	25
17	O	M	U	0	5	10
18	B	F	G	4	10	100
19	O	F	G	7	30	85
20	W	F	G	3	15	100

Just a brief recapitulation of a simple subject (**independent and dependent variables**) that seemed to be easily misunderstood.



Independent variable: Rod and reel.
Dependent variable: Britches.
Population: Goofy.

Now we can relate this to our example above. The independent variable is **Length of Time of Visit (in Minutes)** and the dependent variable is the **Number of Materials Checked Out**.

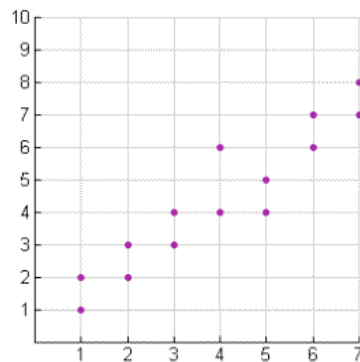
When we look at Slide 6, Variance Accounted For, we are trying to explain the **VARIANCE** in the number of materials checked out. Or stated another way, we are trying to determine the percentage of variance in the **number of items checked** out that can be attributed to the **length of time of the visit**.

We have already calculated Pearson's r to be .42. (Rounding.) And the result of "squaring" the Pearson's r represents the relationship between the two variables and if we convert this squared number into a percentage, this answer will give us the Percentage of Variance (**in the number of items checked out**) that can be Accounted For (**in the length of time of the visit**).

So, $.42 \times .42 = .1764$

If we write ".1764" as a percent we have 17.64%.

We can say that the Length of Time of Visit (independent variable) accounts for 17.64% of the Variance of Number of Materials Checked Out (dependent variable). We also have to say that there is 82.36% (100% - 17.64%) of the variance in the number of items checked out that is not accounted for by length of time the user is in the library.



Working together

Of the two characters above, I do not know which is the independent and dependent variable, but I am sure there is a relatively high positive correlation.

...